White Paper Report

Report ID: 106523

Application Number: HD5157312

Project Director: Trevor Munoz (tmunoz@umd.edu)

Institution: University of Maryland, College Park

Reporting Period: 10/1/2012-9/30/2013

Report Due: 12/31/2013

Date Submitted: 5/8/2014

Project White Paper

Grant # HD-51573-12

ANGLES: A web-based XML Editor

Project Director: Trevor Muñoz
University of Maryland

April 30, 2014

Final Performance Report

ANGLES: A web-based XML Editor was awarded a Level 2 Digital Humanities Start Up Grant by the National Endowment for the Humanities in the amount of $49,929. Led by Assistant Dean for Digital Humanities Research and Associate Director of the Maryland Institute for Technology in the Humanities Trevor Muñoz, ANGLES proposed a bridge between humanities centers who have greater resources to program scholarly software and the scholars who form the core user community for such software through their teaching and research.

The ANGLES project was an experiment in digital humanities tool development. The project proposed a new solution to the adoption gap that has developed between scholars with digital materials and technical developers designing the applications scholars are using in their research. The aim of ANGLES was to go beyond the exploratory, hermeneutic coding of a small group of scholars and to build on the success of the "barnraising" small institute model by combining a model of intensive code development (the "code sprint") with participatory design exercises that testing and feedback from domain experts gathered at internationally-recognized disciplinary conferences. This approach was intended to address the need of a new software product to generate users, provide better technical development by having potential users to explore code as it is being created, and allow the scholarly and developer communities to more clearly elucidate their intellectual, methodological, and technological needs in a structured collaboration process.

As a pilot for this "community roadshow" approach to tool development in the digital humanities, the ANGLES Project worked to develop a web-based editor for working with XML markup through engagement with the large and active community of scholars, teachers, and developers who work with the Text Encoding Initiative.

This final report will review the accomplishment of the project's goals including the release of open source software implementing the web-based XML editor prototype.

Project Activities

Activities from October 1, 2012 – July 31, 2013

Initial technology selection and development
To achieve usable results in the short timeframe of the project, the ANGLES team elected to develop on top of an existing browser-based code editor. The Ace editor (http://ace.c9.io/) is an open-source project under active development. Ace has also been incorporated into larger, commercial projects, most notably the code sharing platform GitHub. Given this level of uptake in the software community, Ace seemed a solid choice for the smaller ANGLES project. The project team, led by Jim Smith, conducted a thorough review of the Ace codebase and began preparing documentation and build tools to allow developers to begin contributing to ANGLES.

Since the main technology under development in the ANGLES Project is a web-based editor for working with XML markup, the TEI Members' Meeting was an important locale for project performance. The 2012 meeting and conference was held November 6-10 at Texas A&M University in College Station, Texas.

The project team submitted a proposal for a pre-conference workshop and a poster presentation; both proposals were accepted. In accordance with the project plan submitted with the original grant application, the ANGLES project funded a group of developers to attend the conference, participate at the "code sprint" workshop, and talk to potential users from the TEI community. The following core developers participated in the first ANGLES workshop:

- Jim Smith, Maryland Institute for Technology in the Humanities (MITH)
- Doug Reside, New York Public Library
- Hugh Cayless, New York University
- Trevor Muñoz, Maryland Institute for Technology in the Humanities (MITH)
- Ondine LeBlanc, Massachusetts Historical Society

The day-and-a-half long code sprint was open to other conference attendees. During this time the core team was joined by five other participants on a drop-in basis. Development tasks completed at this event included:

- Prototype of completely in-browser validation of permitted elements
- Simple API for XML (SAX) parsing of XML document in the editor
- Test integration with Google Drive (as an example storage service)
- Set-up and design of basic demo page

All code is publicly available from the project's GitHub repository:
https://github.com/umd-mith/angles

During the main conference, Trevor Muñoz presented a poster on plans for the project and outcomes of the first code sprint. The poster can be viewed at:
http://dx.doi.org/10.6084/m9.figshare.106812

Additional Development at MITH
After the first code sprint at the TEI Members' Meeting, the MITH team continued some development work to capitalize on insights from the first event and to prepare for subsequent events.

The MITH team made significant changes to the project's codebase to make development easier. Under the supervision of Trevor Muñoz and Jim Smith, the application was re-architected to use Backbone.js (http://backbonejs.org/) a Javascript framework designed to support

full-featured applications in the browser. This technical decision was intended again to align ANGLES with a larger community of open source developers. Writing plug-ins or alternative interfaces for ANGLES will thus not require learning project-specific conventions but merely adopting conventions of the Backbone.js community. Also, this decision would allow ANGLES to easily decouple development from the Ace editor project in the future if this seems desirable. By using the Backbone framework to communicate with Ace, we have created a stable API, which could take a different editor in the future if need be. Additionally, despite progress on a technical implementation of validation against a TEI schema purely in a web browser, the MITH team made a decision to setup up a small web service to provide more full-featured validation using server-side tools such as Jing (http://www.thaiopensource.com/relaxng/jing.html). This web service will complement the more-basic browser-based validation solution for the time being.

Code Sprint at Digital Humanities 2013

The project team organized a smaller, more-focused code sprint at the Digital Humanities 2013 conference, held July 16-19, 2013, in Lincoln, Nebraska. For this second code sprint, the goal was to meet and work with other developers of complex Javascript applications for the digital humanities. The rationale was that these developers were most likely to adopt a complete ANGLES editor into their digital projects; also, the project team hoped that this community might have best practices and insights to share. As the major disciplinary conference, Digital Humanities was also targeted as a crucial site of performance for the ANGLES grant.

The following developers participated in the second ANGLES code sprint:

- Raffaele Viglianti, Maryland Institute for Technology in the Humanities (MITH)
- Jim Smith, Maryland Institute for Technology in the Humanities (MITH)
- Doug Reside, New York Public Library
- Trevor Muñoz, Maryland Institute for Technology in the Humanities (MITH)
- Gregor Middell, Julius-Maximilians-Universität Würzburg
- Wayne Graham, Scholar's Lab, University of Virginia
- David McClure, Scholar's Lab, University of Virginia
- Eric Rochester, Scholar's Lab, University of Virginia
- Sean Pue, Michigan State University
- Wout Dillen, Centre for Manuscript Genetics, Antwerp
- Andrew Hankinson, McGill University

The project team derived some preliminary insights from this event. There was a clear need for better packaging, building tools, and "getting started" documentation, as well as an increased focused on testing existing code that is developed. Iteration was identified as a key concept that would be required by developers with regular validation and review of code.

Activities from August 1, 2013- September 30, 2013

Additional Development at MITH: Improvement of Build Systems and Documentation
Based on feedback from the community received at the second code sprint, the MITH development team focused on activities that would make it easier for developers from other digital humanities projects to incorporate the ANGLES editor as a component in their own projects. Most crucial to reusability was the provision of tools that would reliably allow other developers to download and build source code for the application on their own machines and a test suite by which future developers could evaluate their local modifications to ensure compatibility with the core application.

During this period, the project team incorporated an automated build tool (grunt) into the open source code released as part of the project. This tool provided a series of simple commands by which other developers could automate the downloading and compilation of the various software components involved in a working instance of ANGLES. The use of grunt and similar tools is an increasingly common practice aimed at managing dependencies for complex projects. In addition to the creation of an automated build tool, the project team worked to create a test suite for the code under development. For a project such as ANGLES, focused building a re-usable component that might be used in multiple different project environments, providing ways to isolate and track errors is a crucial inducement for developer adoption. The test suite will allow future developers to introduce modifications and extensions of the core application with confidence. This test suite was integrated with the build tools to allow for automated checking of downloaded components.

Finally, the MITH development team prepared documentation on the architecture of the application, its software dependencies, and the use of the automated build tools and test suite. This documentation has been made available along with the software as a README file in the GitHub repository at: https://github.com/umd-mith/angles

Accomplishments
Despite the challenges encountered in developing software through hackathons or code sprints (discussed below), the ANGLES project succeeded in adapting an existing set of open source software components to create a prototype of an online editor adapted to work with the TEI. The ANGLES editor can be run in a web browser without the need for any external dependencies. In this mode, the checking of the user's XML document uses a simplified form of well-formedness checking and schema validation, which, for instance, can verify that only a selected set of TEI elements should be present in a document, etc. This functionality would support the use of ANGLES in a classroom setting as part of an introductory class session on TEI. By making calls to a small server-based component, the ANGLES editor also allows full-featured well-formedness checking and schema validation (using exactly the same open source components used in most commercial XML-editing software). Pairing the ANGLES editor with a server-side component, such as would be possible if one was incorporating the editor into an

existing online project, supports the use case of crowd-sourced or crowd-contributed TEI data to digital humanities projects.

Audiences

Humanities scholars need tools which specifically address their research needs but, at least as critically, humanities scholars need ways of engaging productively with developers and technologists. The ANGLES project attempted to triangulate these audiences towards a single project goal: the development of the web-based XML editor. The broadest audience for the project is the the open source software community where "ownership" of technologies is broadly distributed among a community of users and responsibility for innovation and maintenance is undertaken by a distributed set of developers, many of whom are themselves also users. There are many examples of successful projects created using variations of this model, including the Text Encoding Initiative (TEI) standard, the Zotero citation management software, the Fedora Digital Repository software, and the Wordpress publishing platform. It was our hope to encourage a robust audience that merged those developers with non-technical humanist specialists who were interested in editing documents with standardized markup. For many humanists—the ability to write code and manage the mechanisms of distributed software collaboration (source code management, ticketing, and similar tools)— are prerequisites outside their abilities. Yet, as they use tools, these humanists often encounter design flaws in the user interface and/or desire additional functionalities. That process generally happens post-beta (or even post-launch). Our goal was to bring those users into the development process to short-circuit the iterative revision process that is often strung out over versions throughout years of releases.

Below is a complete list of those who participated actively with the development process by state and country. Importantly, while the majority of those engaged were from academic institutions, we also were able to incorporate cultural heritage organizations into our development efforts.

Listing of Participants by Country:
USA
  Maryland
  - Jim Smith, Maryland Institute for Technology in the Humanities (MITH)
  - Trevor Muñoz, Maryland Institute for Technology in the Humanities (MITH)
  - Raffaele Viglianti, Maryland Institute for Technology in the Humanities (MITH)

  Massachusetts
  - Ondine LeBlanc, Massachusetts Historical Society

  Michigan State University
  - Sean Pue, Michigan State University

  New York

- Doug Reside, New York Public Library
- Hugh Cayless, New York University

Virginia
- Wayne Graham, Scholar's Lab, University of Virginia
- David McClure, Scholar's Lab, University of Virginia
- Eric Rochester, Scholar's Lab, University of Virginia

Belgium
- Wout Dillen, Centre for Manuscript Genetics, Antwerp

Canada
- Andrew Hankinson, McGill University

Germany
- Gregor Middell, Julius-Maximilians-Universität Würzburg

Evaluation

The ANGLES project did not conduct official evaluations. Instead, it conducted unofficial interviews with leading developers in the digital humanities who reviewed the code base and provided feedback on its quality and consistency. We incorporated those rounds of feedback from the code sprint reviews into the next stages of development thereby following an agile development process as is increasingly common in software engineering.

With regards to the larger place of this project within the digital humanities, the model we attempted for this project met with limited success. Ultimately, the code sprint model attached to conferences did not yield the desired results because major project stakeholders and their developers do not often attend the same conferences. Developers attend their own technical conferences while primary investigators attend the more humanistically-driven events. As such, teams often only send a single individual to conference rather than the full teams that would be best positioned to engage with a software development process. Additionally, the code sprint or hackathon model works better for small applications or data integration or visualization activities (perhaps built on open application programming interfaces, APIs) rather than for the more robust and utilitarian development activities required for building a tool such as an XML editor. As transient events, hackathons need to offer participants either easily-digestible learning opportunities or rapid feedback and gratification. The "meat-and-potatoes" type development of building an editor application does not usually offer either.

It was a result of these dual limitations: the failure of the model to work with the type of tool development we were doing coupled with the limited opportunities to integrate developers and users positioned the project to meet with limited success. After the second code sprint, key developers and staff discussed the failure of the model and it was determined that re-embedding

the project within the traditional digital humanities development cycle would lead to a better quality result. As such, we did not conduct the final code sprint as scheduled.

Continuation of the Project

The ANGLES project currently has no plans to continue development. In part, the decision to continue is a result of developments by commercial XML editors. SyncRO Soft, the company responsible for the commercial XML editing software <oXygen/>, began during the period of this award to release additional versions that addressed the web-based access to their software. With release 15.2 (the most recent) <oXygen/> now functions as a standalone desktop application, a java webstart application, and an eclipse plugin. More importantly, <oXygen/> is now platform agnostic and can be used on mobile devices. Given the tremendous dominance of <oXygen/> within the text encoding and XML editing communities, it does not behoove either the project team or the NEH to continue to devote resources to the project.

Long Term Impact

Given the limitations of the model as well as the development undertaken by <oXygen/>, the long term impact of this project is extremely limited. In fact, we believe that the potential impact is to serve as a case study of the failure of this type of development and to serve as a cautionary tale for future projects who desire a code-sprint based development process that incorporates community engagement.

Grant Products

The project maintains an archived website at:
http://mith.umd.edu/research/project/angles/

All code from the ANGLES project is publicly available from the project's GitHub repository:
https://github.com/umd-mith/angles

A poster on plans for the project and outcomes of the first code sprint was presented at the Text Encoding Initiative Conference 2012. The poster can be viewed at:
http://dx.doi.org/10.6084/m9.figshare.106812